



Externalismus und Selbstkenntnis

Evans' Kompatibilismus

Szanto, Thomas

Published in:
Sprache, Wahrnehmung und Selbst

Publication date:
2016

Document version
Peer-review version

Citation for published version (APA):
Szanto, T. (2016). Externalismus und Selbstkenntnis: Evans' Kompatibilismus. In C. Misselhorn, U. Pompe-Alama, & U. Ramming (Eds.), *Sprache, Wahrnehmung und Selbst: Neue Perspektiven auf Gareth Evans' Philosophie* (pp. 259-282). Mentis Verlag.

Externalismus und Selbstkenntnis:

Evans' Kompatibilismus

Thomas Szanto

1. Einleitung

Kurz nach der Neuauflage der traditionellen Selbstbewusstseins-Debatte in der analytischen Philosophie in den 1960er- bis 80er-Jahren¹ – und erstaunlich wenig davon beeinflusst – ist eine Diskussion darüber entbrannt, ob der semantische oder geisttheoretische Externalismus mit der epistemischen Privilegierung, die Individuen bezüglich ihrer eigenen mentalen Zustände genießen, kompatibel sei.²

In einem ersten Vorgriff lässt sich das Dilemma, um das sich Kompatibilisten und Inkompatibilisten in Bezug auf Selbstkenntnis ringen, folgendermaßen konstruieren: Wenn der Externalismus die richtige Theorie der Individuation mentaler Zustände ist, dann, so scheint es, können Subjekte keine *epistemisch privilegierte Selbstkenntnis* (EPS) haben; umgekehrt, müssen wir, sofern wir an der Idee von EPS festhalten wollen, den Externalismus verabschieden. Beide Optionen, sind aber nicht sehr attraktiv, und für manche sogar inakzeptabel – daher das Dilemma.

Die Wege und Umwege dieser äußerst komplexen Diskussion³ haben zu der paradoxen Situation geführt, dass es gerade die leidenschaftlichsten Vertreter des geisttheoretischen Externalismus (und nicht, wie man vermuten würde, vielmehr die Internalisten), die sich gezwungen sehen, angesichts der inkompatibilistischen Argumentationsstrategien immer ausgefeiltere externalistische Theorien zu entwickeln, um EPS – dem sie ebenso wie die Inkompatibi-

¹ Gemeint sind hier insbesondere Strawson 1959; Castañeda 1966; Shoemaker 1968; Rorty 1970; Lewis 1979; Perry 1979; Chisholm 1981; Anscombe 1981 und eben Evans 1982.

² Als diskursbegründend können die Arbeiten von Burge (1979; 1988) und Davidson (1984b; 1987; 1988a; 1989). Daran anknüpfend reihen sich u. a. Heil 1988; Bilgrami 1992; Brueckner 1992, 2007; Warfield 1992; Falvey & Owens 1994; Dretske 1995 (insbes. 64ff.); Macdonald 1998; Miller 1997; McLaughlin & Tye 1998; Sawyer 1998; Falvey 2000 und Jacob 2004 – die mit z. T. sehr unterschiedlichen Konzepten von Selbstkenntnis, aber mehr oder weniger eindeutig dem kompatibilistischen Lager zuzurechnen sind – und ihre wichtigsten Widersacher McKinsey 1991, 2007; Brown 1995, 1999, 2004 und Boghossian 1989, 1997. Konzise Überblicke liefern Ludlow & Martin 1998; Wright et al. 1998 und Nuccetelli 2003 bzw. Rowlands 2003, Kap. 8 und Brown 2007, und Kallestrup 2011 und am rezentesten Goldberg 2015.

³ Ich habe mich mit dieser Debatte weit ausführlicher und aus einer transzendentalphänomenologischen Perspektive kritisch in Szanto 2012 (399-462) auseinandergesetzt.

listen eine gewisse intuitive Anziehungskraft einräumen – irgendwie unterbringen zu können.

Evans ist kurioserweise in dieser Debatte bislang ein praktisch unbeschriebenes Blatt, obwohl seine Theorie der Selbstkenntnis enorm einflussreich ist. Angesichts dieses Desiderats möchte ich im Folgenden zeigen, dass Evans, auch wenn er keine explizite und vielleicht auch nicht, wie Burge oder Davidson, die eleganteste, so doch eine der überzeugendsten Lösungsvorschläge bereithält. Ich werde mit Evans dafür argumentieren, dass sowohl die externalistische Bedrohung für EPS als auch die vermeintliche einzige Alternative, den Externalismus aufzugeben, sich als unbegründet erweisen.

Ich werde dabei folgendermaßen vorgehen: In Abschnitt 2, werde ich die das Dilemma und die unterschiedlichen Argumente für den Inkompatibilismus diskutieren. In Abschnitt 3 werde ich Evans spezifisch informationstheoretische, radikale Version des Externalismus darlegen und ein Argument konstruieren, das Evans' vermeintlichen Inkompatibilismus demonstrieren soll. Anschließend werde ich Evans' Lösungsstrategie vorstellen (Abschnitt 4). Ich werde mich dabei insbesondere auf Evans' komplexe Theorie der Selbstidentifikation mit Blick auf unser gegenständliches Problem konzentrieren (4.1), um dann seinen alternativen Kompatibilismus skizzieren (4.2). Schließlich werde ich auf die wichtigsten Einwände gegen Evans antworten (Abschnitt 5.), um dann den Schluß zu ziehen, dass, selbst wenn sie zuträfen, dies das kompatibilistische Kernargument von Evans nicht ernsthaft gefährdet.

2. Externalismus und Selbstkenntnis: Das *Prima facie*-Dilemma

Die Kompatibilität des privilegierten Zugangs zu den eigenen mentalen Zuständen mit dem Externalismus hängt freilich zum einen davon ab, was der Externalismus speziell in Bezug auf Selbstkenntnis genau besagt, und zum anderen davon, was man unter epistemischer Privilegiertheit eigentlich versteht. Zunächst zum ersten Punkt.

Der Externalismus in Bezug auf die Individuation mentaler Zustände wird in der betreffenden Auseinandersetzung üblicherweise unter dem Titel *Anti-Individualismus* (AI) verhandelt. Diese Bezeichnung geht auf Burge zurück, und streicht den (methodologisch) solipsistischen Aspekt des Internalismus heraus (Burge 1979; 1986; 1988). AI besagt grob, dass die Individuation der mentalen Zustände eines Individuums bzw. deren Inhalt *notwendig* von jenen (kausalen und sozialen) Relationen abhängt, die das betreffende Individuum zu seiner physikalischen Umwelt und anderen Individuen einer Sprachgemeinschaft einnimmt. Das Problem, das der Anti-Individualismus speziell für EPS zu stellen scheint, resultiert aus einer epistemisch modifizierten Version des semantischen und geisttheoretischen Externalismus⁴, die man so formulieren kann:

⁴ Für weithin akzeptierte, konzise Formulierungen des geisttheoretischen und semantischen Externalismus, vgl. McGinn 1989 und Rowlands 2003.

- (AI) Die externen Faktoren, die die mentalen Zustände eines Subjekts S individuieren, sind gänzlich (logisch-konzeptuell und/oder metaphysisch) unabhängig vom introspektiven *Wissen*, das S von seinen mentalen Zuständen und deren Gehalt hat, wobei das Wissen, das S von seinen mentalen Zuständen haben soll, *epistemisch privilegiertes* (und evtl. gerechtfertigtes) Wissen sein soll.

Angesichts von AI wird nun das Problem, um das sich die Auseinandersetzung zwischen Kompatibilisten und Inkompatibilisten zentriert – wohlgermerkt aus der Perspektive der Inkompatibilisten –, üblicherweise als ein klassisches Dilemma konstruiert (vgl. Heil 1988): Entweder wir nehmen uns die weithin akzeptierten Lehren des Anti-Individualismus zu Herzen, oder aber wir pochen auf EPS, was im gegebenen Fall zugleich hieße, dass EPS über jede externe Beschränkung seitens der Welt oder anderer Subjekte erhaben ist. Beides jedenfalls könnten wir nicht haben. So weit, so gut; aus den beiden, sich vermeintlich ausschließenden Optionen allein ergibt sich aber noch lange kein *Dilemma*. Das eigentliche Dilemma des Inkompatibilismus entsteht nun aus der Annahme, dass, wofür wir uns immer entscheiden, inakzeptable bzw. kontra-intuitive Konsequenzen folgten, oder wir uns unauflösliche Probleme einhandelten.

Das erste Horn des Dilemmas besteht darin, dass AI vorausgesetzt es rätselhaft erscheint, dass wir *meistens bestens* darüber Bescheid wissen, *wovon* unsere Gedanken handeln und *wie* wir über etwas Bestimmtes denken – unerachtet dessen, dass unsere Gedanken *oft* nicht mit dem übereinstimmen, wie die Welt wirklich beschaffen ist, unerachtet auch dessen, dass wir nicht immer *wissen*, wie der bestimmte Gegenstand unserer Gedanken tatsächlich beschaffen ist und ob der Gegenstand oder Sachverhalt, an den wir denken, vorliegt oder nicht – und uns gar, wenn auch in seltenen Fällen, darüber täuschen können, *wie* wir an etwas denken (z. B. ob wir, was auch immer tatsächlich der Fall sein mag, nun etwas fürchten, lieben, oder verabscheuen). Zudem, so ein weiterer oft vorgebrachter Aspekt, bliebe es schwer verständlich, wie eine mögliche Ignoranz der *Individuationsbedingungen* unserer Gedanken, die wir ja tatsächlich oft nicht kennen oder deren wir zumeist nicht gewahr sind, sich in keiner Weise auf die Kenntnis unserer Gedanken selbst auswirken sollte. Schließlich müssten wir, wenn es zutrifft, dass unsere Gedanken durch Faktoren bestimmt werden, die ihnen äußerlich sind, diese Faktoren begutachten, um festzustellen, welche Gedanken wir worüber haben – eine Prozedur der Informationsbeschaffung, die jedoch völlig unseren normalen kognitiven Praktiken zuwiderläuft.

Das zweite Horn des Dilemmas besteht darin, dass die Annahme einer epistemischen Privilegierung uns zu einer Art Privation zu führen scheint, zu einer Loslösung von der Welt und den anderen Subjekten. Wenn wir auf unser Privileg bestehen, *einzig* und allein und auch noch *verlässlich* darüber informiert zu sein, was und wie wir denken, unabhängig davon, was andere denken und ob die Welt mitmacht, steht jeder von uns am Ende allein (mit seinen eigenen Gedanken) da.

Mit anderen Worten: Wenn man den epistemischen Zugang, der mit EPS vorliegt, zu stark externalistisch einschränkt, entsteht das Problem, dass der Zugang zu unseren eigenen Gedanken gegen all unsere Intuitionen und unsere normale (Selbst-)Erfahrung *potenziell permanent* unterminiert wird. Wenn man dagegen den Zugang rein internalistisch fasst, bleibt dem Skeptiker jederzeit die Hintertür offen, und es wird in einem ebenfalls inakzeptablen Maße fraglich, ob unsere Gedanken überhaupt von irgendetwas, außer von sich selbst, handeln. Burge – neben Putnam nicht nur der wohl wichtigste Anti-Individualist, sondern neben Davidson auch der prominenteste Vertreter des Kompatibilismus – hat das treffend formuliert (vgl. auch Davidson 1987, 663):

Our problem is that of understanding how one can know some of our mental events in a direct, nonempirical manner, when those events depend for their identities on our relations to the environment. A person need not investigate the environment to know what his thoughts are. A person has to investigate the environment to know what the environment is like. Does this not indicate that the mental events are what they are independently of the environment? (Burge 1988, 650)⁵

Zu beachten ist dabei, dass die Kompatibilität von Anti-Individualismus und der Möglichkeit von EPS nicht notwendig davon abhängt, ob EPS als *wahres* oder *gerechtfertigtes propositionales* Wissen qualifiziert wird und auch nicht davon, ob die betreffenden Gedanken erfolgreich referieren oder nicht. Kompatibilisten können behaupten, dass Selbstkenntnis unabhängig davon *epistemisch privilegiert* ist, ob der propositionale Gehalt eines mentalen Zustandes erfüllt ist bzw. ob der semantische Gehalt der jeweiligen mentalen Zustände erfolgreich referiert. Demzufolge wird die Kompatibilitätsthese von möglichen skeptischen Szenarien nicht tangiert.

Genau das bezweifeln nun Inkompatibilisten. Inkompatibilisten halten AI genau insofern für problematisch als ihrer Ansicht nach skeptische oder externalistische Zwillingserde-Szenarien (à la Putnam 1975) entweder den Aspekt der *epistemischen* Privilegierung von EPS unterminieren, oder aber die Plausibilität von AI (vgl. Brueckner 1992; Falvey & Owens 1994). Wie ist das zu verstehen?

Inkompatibilisten konstruieren EPS als eine Form *apriorischen* Wissens und sprechen dabei üblicherweise vom *Prinzip des privilegierten Zugangs*. Demnach haben Subjekte apriorisches Wissen von ihren eigenen mentalen Zuständen und Gedanken. ‚Apriorisch‘ kennzeichnet hier ein Wissen, das unabhängig von empirischen Untersuchungen erworben wird. Empirische Untersuchungen sind wiederum all jene Untersuchungen, die die physikalische Umwelt des jeweiligen Subjekts betreffen. Wir können das Prinzip – das Evans, wie wir noch sehen werden, ablehnen wird – auch so formulieren:

(PEPZ) Ein Subjekt S hat epistemisch privilegierten Zugang zu seinen eigenen mentalen Zuständen Z, wenn S's introspektives Wissen über Z aprio-

⁵ Burges kongenialen kompatibilistischen Lösungsvorschlag habe ich, neben jenen von Davidson und McDowell, in Szanto 2012 (Kap. III. 3.1, 3.2) diskutiert.

risches Wissen ist, d. i. unabhängig von Beobachtung oder sonstigen empirischen Untersuchungen erworben wird (im Gegensatz zu S's Wissen über seine physikalische Umwelt).

Die Standardargumente der Inkompatibilisten, die PEPZ voraussetzen, gehen auf die Arbeiten von McKinsey (1991) und Boghossian (1989, 1997) zurück und haben die Form einer Reductio ad absurdum.⁶ Die Reductio soll zeigen, dass der Kompatibilismus notwendig zu der unannehmbaren These führt, dass ein Subjekt gleichsam inferenziell – und das heißt hier, via die autoritative, apriorische Struktur seiner Selbstkenntnis, also via PEPZ – *entweder* auf die externen Individuationskriterien seiner Gedanken *oder* auf die Existenz der konkreten Individuationsbedingungen oder auf die Existenz der Gegenstände, von denen seine Gedanken handeln, schließen könnte. Nach McKinsey (1991) vertritt, genauer, der Anti-Individualist à la Burge notwendig drei Thesen, die miteinander nicht kompatibel sind:

- (1) Wenn ein Subjekt S EPS hat, dann weiß S a priori, dass S denkt, dass *p*.
- (2) Die Proposition P, „dass S denkt, dass *p*“, hängt (logisch/konzeptuell) notwendig von E ab, wobei E irgendeine externe Tatsache, Eigenschaft, Gegenstand oder Proposition ist, jedenfalls genau dasjenige, was P (typ- oder token-)individuiert. [Folgt aus der semantischen Version von AI.]
- (3) Kein Subjekt (also auch nicht S) hat von E apriorische Kenntnis; Kenntnis von E kann ein Subjekt vielmehr nur durch empirische Untersuchung seiner Umwelt erwerben.

Zwischen (1) und (3) herrscht in der Tat eine unauflösbare Spannung. Entweder man gibt die These auf, dass ein Subjekt EPS-begabt ist und also apriorische Kenntnis von seinen propositionalen Einstellungen hat, oder man gibt die nach McKinsey zentralen externalistischen Thesen (2) und (3) auf. Solange diese Spannung nicht aufgelöst werden kann, scheint jedenfalls der Kompatibilismus keine sinnvolle Theorie zu sein.⁷

Das Standard-Reductio-Argument gegen den Kompatibilismus lässt sich auch ohne Rückgriff auf propositionale Einstellungen und ontologisch noch prägnanter folgendermaßen darstellen (vgl. McKinsey 1991):

- (1) Die Existenz eines mentalen Zustandes, Gedankens etc. eines Subjekts S impliziert *logisch* oder *konzeptuell* die Existenz von etwas (Gegenstand, Eigenschaft, Proposition etc.), das S extern ist.
- (2) Wenn ein Subjekt S a priori davon Kenntnis hätte, dass es sich in einem bestimmten mentalen Zustand befindet, dann würde S a priori

⁶ Eine etwas andere inkompatibilistische Argumentation, die nicht auf einer Reductio basiert, findet sich in Boghossian 1989; vgl. dazu kritisch: Warfield 1992 und siehe auch Ludlow 1995.

⁷ Vgl. die einflussreiche Kritik bei Brueckner 1992 und die Auseinandersetzung zwischen Brueckner 2007 und McKinsey 2007.

- wissen, dass die externe Welt existiert. [Folgt aus der logisch-konzeptuellen Implikation, die in (1) formuliert ist.]
- (3) Es ist offensichtlich, dass niemand a priori Kenntnis von der Außenwelt hat.
 - (4) Niemand kennt a priori seine eigenen okkurrenten mentalen Zustände.

Nimmt man die Reduktion ernst, gibt es aber immer noch zwei Optionen. Entweder man macht bei (3) halt, dann folgt, dass der Anti-Individualismus nicht die richtige Theorie der Individuation des Gehalts (bestimmter) mentaler Zustände ist; oder aber man geht einen Schritt weiter zu (4). Wofür man auch optiert, das Argument zeigt jedenfalls, dass EPS und der Anti-Individualismus nicht kompatibel sind.

Doch wie plausibel ist nun die Reduktion selbst? Trifft die Absurdität, die mit solchen Reductio-Argumenten aufzeigt werden, überhaupt die Plausibilität der Annahme von EPS? Welche implizite Zusatzannahme liegt solchen Argumenten zugrunde? Und zielen sie nicht am eigentlichen Phänomen von EPS vorbei? Soweit ich sehe, ist genau dies der Fall.

Inkompatibilistische Reductio-Argumente operieren mit einer Art Verdoppelung des (propositionalen) Wissens von Subjekten. Demnach müssten Subjekte entweder a.) (*pace* Burge) selbst die Individuationskriterien und/oder die jeweiligen Individuationsbedingungen kennen, um die eigenen Gedanken zu kennen, oder b.) (was zunächst noch absurder erscheint, aber eine schwächere These ist) zumindest müssten genau diejenigen Subjekte, die apriorische Kenntnis von (der Geltung von) anti-individualistischen Argumenten haben, a priori wissen, dass, wenn sie einen bestimmten Gedanken über etwas haben, und sie introspektiven Zugang zu diesem Gedanken haben, jene Objekte, Eigenschaften, Sachverhalte, über die ihre Gedanken handeln, existieren bzw. bestehen.⁸ Beide Forderungen, die – aus der Perspektive des Inkompatibilisten – die Kompatibilität von Anti-Individualismus und EPS garantieren sollen, scheinen jedenfalls eine äußerst komplizierte epistemische Prozedur nach sich zu ziehen. Beide Forderungen ließen sich in der folgenden Zusatzannahme, die man die *Kompatibilitätsthese* (KT) nennen könnte, zusammenfassen:

- (KT) Damit der Kompatibilismus möglich ist, muss ein (externalistisch konstruiertes, EPS-begabtes) Subjekt S a priori wissen (können), dass jede Proposition des Typs P, die S hat, notwendig (konzeptuell oder logisch) abhängig ist von einer Proposition P*, die besagt, dass Propositionen des Typs P von der Existenz von S externer Tatsachen abhängig sind (bzw. durch die Relation, die S zu diesen Tatsachen hat, individuiert werden).⁹

⁸ Boghossian 1989, 73 und Brown 1995, 149 weisen explizit darauf hin, dass die erste Alternative (a.) kein Prinzip des Inkompatibilismus sein muss und auch nicht sein soll.

⁹ Die prägnanteste Formulierung von KT findet sich in Brown 1995, die die These selbst gegen Kompatibilisten wendet.

Einmal abgesehen davon, dass es äußerst unplausibel ist, dass Subjekte im Falle *jeder* Proposition die Kenntnis, die KT voraussetzt, haben, und auch davon, dass es offensichtlich falsch ist, dass *alle* Subjekte (etwa jene, die keine philosophischen Abhandlungen über den Externalismus gelesen haben) solche *apriorische* Kenntnis haben – der Grund, warum KT am Problem der Selbstkenntnis vorbeizieht, ist ein anderer. Der Grund ist, dass EPS nichts damit zu tun hat, ob Subjekte apriorische oder aber empirische Kenntnis von ihren Gedanken oder der (Existenz der) konkreten Individuationsbedingungen ihrer Gedanken und deren Inhalte haben. Dies möchte ich nun mit Evans zeigen.

3. Evans Anti-Individualismus

Für jemanden, der auch nur ein wenig mit Evans' Werk vertraut ist, mag es zunächst scheinen, dass sich gerade für oder mit Evans das inkompatibilistische Dilemma noch brisanter stellt. Und in der Tat ist kaum jemand radikaler in seinem Externalismus in Bezug auf Selbstkenntnis als Evans (und Evans' Wittgenstein). Wie ich jedoch im Folgenden argumentieren werde, sind für Evans zwei scheinbar inkompatible Thesen durchaus vereinbar: nämlich seine berühmt-berüchtigte, radikal-externalistische *Transparenzthese bezüglich Selbstkenntnis* (TS), auf der einen Seite, und seine These bezüglich der *Identifikations-Unabhängigkeit von Selbstkenntnis* bzw. seine ebenfalls wohlbekannte These von der *Immunität gegen Fehlidentifizierung* (*Immunity to Error through Misidentification*) von Ich-Gedanken (IF) auf der anderen Seite.

TS zufolge wissen wir, was wir denken, nicht, weil wir uns in irgendwelchen mentalen Zuständen befinden, sondern weil und sofern sich unsere Gedanken auf externe Entitäten *beziehen*. Von diesen beziehen sich unsere Gedanken ihren kognitiven Gehalt (Information), und wir mithin (empirische) Informationen von der Welt. Wenn sich unsere Gedanken auf nichts beziehen, oder wir nicht wissen, worauf sie sich beziehen, weil wir keine Informationsquelle oder falsche Informationen haben, gibt es rein gar nichts, was wir sonst konsultieren könnten, um zu wissen, an welchen bestimmten Gegenstand wir denken; mithin haben wir gar keine Gedanken. Diese These scheint auf den ersten Blick jedoch direkt der These von einer bestimmten Irrtumsresistenz bezüglich unserer eigenen Gedanken zuwiderzulaufen. Dieser These, IF, zufolge, ist nämlich die Semantik von selbst-referenziellen, sogenannten Ich-Gedanken (*I-thoughts*) so beschaffen, dass sie eine grundlegende Selbstidentifikation voraussetzt, die beobachtungs- und identifikationsunabhängig und immun gegen Fehlidentifizierung ist. Dies wiederum soll eine Art Wissen (Selbstkenntnis)¹⁰ ermöglichen, welches weitgehend – wenn auch nicht vollständig –

¹⁰ Ob identifizierungsfreie (Ich-)Gedanken nun tatsächlich Instanzen von *Wissen* (oder auch von *Selbstkenntnis* im prägnanten Sinne) sind, ist bei Evans nicht ganz klar, daher auch der Zusatz *„eine Art Wissen“*; ich werde die betreffenden epistemologischen Differenzierung hier jedoch außer Acht lassen, da sie für mein Kernargument keine wesentliche Rolle spielen. Ich werde jedenfalls IF bzw. die semantisch-epistemische Struktur Ich-Gedanken weiter unten (Abschnitt 4.1) genau diskutieren.

unabhängig ist von den Objekten unserer Gedanken und der Kenntnis ihrer externen Individuationsfaktoren.

Doch um zu verstehen, inwiefern sich für Evans aufgrund dieser beiden Annahmen, TS und IF, überhaupt das inkompatibilistische Dilemma stellen würde, müssen wir als zunächst eine genau Vorstellung von Evans' externalistischen Prämissen gewinnen. Das ist umso wichtiger, als sein semantischer und geisttheoretischer Externalismus in der Mainstream Internalismus/Externalismus-Debatte bisher praktisch keine Beachtung fand, obwohl sie eine äußerst interessante, informationstheoretisch angereicherte Version des Externalismus darstellt (vgl. Rammings Beitrag in diesem Band). Anschließend müssen wir uns die zentralen semantischen und epistemologischen Prinzipien, die Evans' spezifischer Externalismus voraussetzt, und die von den semantischen Standardannahmen einer kausaler Referenztheorie der Art von Kripke (1980) und Putnam (1975) zum Teil gravierend abweichen, erläutern.

Was Evans' Externalismus betrifft, kann kaum genug betont werden, dass er nicht nur eine sogenannte neo-russellsche, sondern auch eine kausal-externalistische Theorie der Referenz vertritt (eine Theorie übrigens, die Evans bereits sehr früh entwickelte; vgl. Evans 1973, 1980, und 1982 [im Folgenden VR für *Varieties of Reference*], 76-79). Dies zu betonen scheint umso eher angebracht, als es bei der Komplexität der betreffenden Referenz-Theorie leicht übersehen oder missverstanden werden könnte, zumal Kripke (1980) bekanntlich die Kausaltheorie der Referenz zum Teil direkt gegen Russells Beschreibungstheorie von Eigennamen und Kennzeichnungen lanciert (und zum anderen Teil, wie stärker noch Putnam (1975), gegen Freges Theorie intensionalen Sinns, beides Theorien, die Evans ja zu integrieren versucht; vgl. insbes. VR, Kap. 1; vgl. auch Kripke 2008). Ferner ist Evans' Externalismus umso interessanter als sie nicht nur eine spezifisch informationstheoretische Variation der Kausal-Theorie der Referenz à la Kripke, Putnam & Co. ist; er stellt auch, weniger aufgrund seiner kausaltheoretischen, als vielmehr gerade aufgrund seiner neo-russellschen Anleihen eine besonders radikale Form des Externalismus dar. Denn Evans' neo-russellsche Theorie von referenziellen Termen und Gedanken (*Russellian terms/thoughts*) hat zur Folge, dass, sofern extra-mentalen Gegenstände nicht existieren, die betreffenden Gedanken nicht nur nicht veridisch sind oder nicht referieren, sondern die Gedanken selbst nicht existieren (VR, 43-79, 173); das aber ist eine besonders extreme Ablehnung jeglicher Form von geisttheoretischer oder semantischer Internalismus. Evans' Externalismus (EE) lässt sich nun folgendermaßen charakterisieren:

- (EE) Die Konstitutions- und Individuationsbedingungen einer (russellschen) referenziellen Bezugnahme eines Subjekts S sind abhängig von der Existenz eines demonstrativ und raum-zeitlich lokalisierbaren Gegenstandes x in der Umwelt von S (d.i., auf der egozentrischen kognitiven Karte von S) und einer Relation R zwischen x und S [$R(S, x)$], wobei R einen epistemischen und referenziellen Kontakt zwischen S und x herstellt, derart, dass S Informationen über x erhält.

Was EE insbesondere von anderen externalistischen Theorien der Referenz unterscheidet, ist die darin formulierte epistemische These bezüglich eines spezifisch ‚informationellen Links‘, der zwischen dem Subjekt und Referenzgegenstandes besteht. Ein informationeller Link hat im Wesentlichen drei Funktionen: a) Die aktuelle oder auch vergangene (und mithin vom Subjekt erinnerte) Informationen, die ein solcher Link liefert, steuert die (perzeptuelle oder erinnerte) Gegenstands-Konzeptionen des Subjekts; b.) das Subjekt bleibt, dank eines solchen Links und nur wenn ein solcher gegeben ist, mit dem Gegenstand in direktem epistemischen Kontakt, sodass es unmittelbar disponiert ist, seine Gegenstands-Konzeption entsprechend zu ändern, wenn sich die Informationen (in Zukunft) ändern; c.) schließlich ermöglicht er dem Subjekt, das Objekt (oder sich selbst) im egozentrischen Raum zu lokalisieren (vgl. VR, 174, 244f.).

Doch, sosehr Informationslinks für die Etablierung einer referenziellen *Relation* unerlässlich sind, sie reichen für Evans gewiß nicht aus um eine referenzielle *Identifikation* zu ermöglichen. Hierfür greift Evans auf Einsichten Russells zurück, und damit kommen wir zum nächsten Punkt, den semantischen *und* epistemologischen Prinzipien, die Evans’ spezifischer Externalismus voraussetzt.

Evans’ übernimmt, reformuliert und ergänzt insbesondere drei Einsichten von Russells Theorie der Kennzeichnungen, wobei er noch einmal – und in Einklang mit seiner informationstheoretischen Wendung des Externalismus – die epistemologische Dimension einer Theorie der Referenz hervorstreicht: In aller Prägnanz und Evans’ Terminologie folgend haben wir also zunächst *Russell’s Principle*, wonach jedes Urteil eines Subjekts S über einen Gegenstand x impliziert, dass S *Kenntnis* von x hat (VR, 64ff., 74, 91-100). Was es aber heißt, von einem Gegenstand Kenntnis zu haben (jenseits der Tatsache wohlgemerkt, dass man dazu in einer Informationsrelation steht), expliziert Evans mithilfe der sog. Allgemeinklausel (*Generality Constraint*). Demzufolge hat S Kenntnis von x genau dann, wenn S einen Begriff davon hat, dass ein bestimmter Gegenstand a die Eigenschaft F hat (F(a)). Einen solchen Begriff zu haben ist wiederum gleichbedeutend damit, dass S weiß, was es heißt, dass eine Proposition der Form F(a) wahr ist, oder mit anderen Worten, dass S die konzeptuelle Fähigkeit hat, beliebige Gedanken der Form F(a) oder G(a), oder F(b) oder G(b) zu bilden (VR, 100-105). Dies aber beruht auf einer weiteren epistemischen Leistung, nämlich auf die Fähigkeit, eine grundlegende Vorstellung (*Fundamental Idea*) von einem Gegenstand zu haben. Demnach hängt S Kenntnis in der Allgemeinklausel davon ab, ob S eine grundlegende Vorstellung δ von x hat, wobei δ durch die grundlegenden Individuationsbedingungen von x festgelegt ist. Diese grundlegenden Individuationsbedingungen sind all jene Faktoren, welche den grundlegenden Unterscheidungsgrund (*fundamental ground of difference*; VR, 107) eines Objekts im Vergleich zu (allen) anderen Objekten markieren (VR, 105-112).

Wir sind nun gerüstet, mit Evans’ ein mögliches inkompatibilistisches Argument konstruieren. Zu beachten ist, dass wir bei Evans selbst kein solches Argument finden, und dass dies, wie wir noch sehen werden, durchaus nicht zufällig, sondern aus systematischen Gründen der Fall ist. Mehr noch, wie ich

zeigen werde, lehnt Evans bestimmte Voraussetzungen des Arguments, die inkompatibilistische Konklusion erst ermöglichen, geradeheraus ab. Das Argument könnte jedenfalls so aussehen:

- (1) Wenn S einen (russellschen) Gedanken über G hat, dann gab oder gibt es ein G und einen kausal-informationalen Link zwischen S und G (auf den S's G -Gedanke, zusammen mit einer grundlegenden Vorstellung δ der Form ' δ ist G ' beruht).
- (2) S hat einen (russellschen) Gedanken über einen Gegenstand G .

Nun kann S durch bloße Introspektion seines G -Gedankens, welches S Wissen über (1) und (2) liefert, zusammen mit einer Modus-ponens-Inferenz – und das heißt eben a priori – auf (3) schließen:

- (3) Es gab/gibt ein G und einen kausal-informationalen Link zwischen S und G .

Das Argument mündet allerdings in das inkompatibilistische Dilemma, da ja (3) plausibler Weise keine Instanz apriorischen Wissens sein kann. Im nächsten Abschnitt möchte ich nun zeigen, warum (3), bei der richtigen Konzeption von Ich-Gedanken und Selbstkenntnis, für Evans jedoch gar nicht folgt.

4. Evans' Kompatibilismus

4.1 Ich-Gedanken und Immunität gegen Fehlidentifizierung

Um Evans Kompatibilismus richtig ermessen zu können, müssen wir eine klare Vorstellung von den semantischen und epistemischen Besonderheiten von Ich-Gedanken bzw. der obigen These IF gewinnen. Diese weisen zwar einige der Kernaspekte der referenziellen Struktur von russellschen Gedanken im Allgemeinen auf, haben aber darüber hinaus einige epistemische Besonderheiten, die mit der besonderen Weise von Identifikation in *Selbstidentifikationen* und der Identifikation von *Personen* zu tun haben. Allgemein besehen sind Ich-Gedanken selbst-referenzielle und selbstbewusste Gedanken der Form 'Ich denke/glaube/meine, etc., dass ich ϕ ', wobei ϕ nicht nur für irgendeinen mentalen oder psychologischen Ausdruck (z. B. 'denken', 'habe Schmerzen'), sondern auch etwa für demonstrative Referenzen oder beliebige andere Beschreibungen, die *mich* im ϕ -Ausdruck selbst einschließen (z. B. 'Ich denke, dass ich bei der nächsten roten Ampel stehenbleiben sollte'), oder sogar weitere propositionale Einstellungen stehen kann (wie etwa 'Ich glaube, dass ich *gestern fälschlicherweise daran gezweifelt habe, dass p*'). Genauer nun weisen Ich-Gedanken zunächst folgende Komponenten auf (VR, 205-209): Erstens haben sie, wie alle russellschen Gedanken, eine Informationskomponente. Diese ist dafür verantwortlich, dass das betreffende Subjekt einen epistemisch-kausalen Kontakt mit nicht-mentalen, extern indivi-

duierten, aber auch selbstbezüglichen, z. B. auf die eigene Person, Handlung oder den eigenen Körper bezogene Informationen hat. Ebenfalls analog zu russellschen Gedanken basieren Ich-Gedanken ferner auf einer grundlegenden ‚Ich‘-Vorstellung, die aber anders als gegenständliche Identifikationen ein Identifikations-*unabhängiges* Wissen, oder eben eine *Selbst*identifikation, zum Resultat hat. Was ein Subjekt in einer solchen Selbstidentifikation nun repräsentational, aber unmittelbar, d. i., ohne sich selbst oder die Umgebung beobachten zu müssen, herausgreift, ist seine Lokalisierung in einem egozentrischen Raum. Dieses Identifikations-unabhängige Wissen ist aber nicht die einzige Art der Individuation des eigenen Ich in Ich-Gedanken. Vielmehr teilen, drittens, Ich-Gedanken mit allen anderen russellschen Gedanken eine bestimmte Form der Allgemeinklausel. (Dies ein zentraler Aspekt von Ich-Gedanken ist, der in der breiten Literatur über Evans berühmte Immunität gegen Fehlidentifizierung gern unterschlagen wird). Demgemäß beruhen Ich-Gedanken auf einer fundamentalen Identifikation einer beliebigen *Person* gemäß der Allgemeinklausel, die wiederum von der Fähigkeit abhängig ist, dass das betreffende Subjekt sich selbst als eine körperliche Person auf einer objektiven, d. i., einer *nicht*-egozentrischen, sondern allozentrischen, raum-zeitlichen Karte verorten kann. Schließlich weisen Ich-Gedanken auch wesentlich eine Handlungskomponente auf, die in der genuin praktischen Fähigkeit besteht, Ich-Gedanken in Handlungen manifestieren zu können und entsprechend auch in Handlungen zu realisieren (vgl. auch Bermúdez 2005, 186ff.)

Die weitere Charakteristik von Ich-Gedanken, die den entscheidenden Schritt hin zu Evans' Kompatibilismus bereitet, ist nun seine berühmte, aber meistens nicht hinreichend genau dargestellte Eigenschaft der erwähnten *Immunität gegen Fehlidentifizierung* (IF) von Ich-Gedanken (VR, 179ff., 237f.). Demnach ist ein Ich-Gedanke der Form ‚Ich bin ϕ ‘ immun gegen Fehlidentifizierung (durch das Subjekt des Ich-Gedankens), wenn die Information, dass ϕ instanziiert ist, nicht auf einer Identifikation eines Paares von Propositionen der Form [δ ist ϕ ‘ und δ = Ich] beruht, sondern vielmehr die Information, mit dem die Konzeption von ‚ ϕ ‘ assoziiert ist, unmittelbares Wissen zu t darüber liefert, dass eben *Ich* (und niemand anderer) ϕ bin; dies wiederum ist genau dann der Fall, wenn die Instanzierung von ϕ zu t Teil des Individuationsgrundes vom Ich-Subjekt ist (als diese ϕ -beschaffene *Person* zu t), sodass die Frage zu t ‚Es gibt eine Person, die ϕ ist, aber bin *ich* diese Person?‘ keinen Sinn macht. Warum macht diese Frage nun keinen Sinn? Nun, deshalb, weil S keine weiteren Informationen – die über die Information, die mit der Konzeption von ϕ gleichsam automatisch mitgeliefert wird und die mir anzeigt, dass *ich* ϕ bin – konsultieren muss, oder auch nur sinnvollerweise konsultieren *kann*, um die Frage zu beantworten.

Genau genommen ist IF keine epistemische, sondern eine referenzielle Besonderheit, nämlich von Urteilen, die auf Identifikations-unabhängige Typen von Gedanken beruhen (VR, 219). Folgende Typen von Propositionen qualifizieren sich dafür: demonstrative Identifikationen wie Dies und Hier (VR, 179ff.), und verschiedene Formen von reflexiven Ich-Gedanken, nämlich solche, die auf mentale (VR, 224ff.), oder körperliche Selbstzuschreibungen basie-

ren (VR, 220ff.), oder jene Selbstzuschreibungen, die auf Erinnerung beruhen (VR, 235ff.).

Nun tritt anstelle des Prinzips epistemisch privilegierter Selbstkenntnis (EPS), auf die die Kompatibilismus-Debatte beruht, bei Evans IF. IF selbst ist eine direkte Konsequenz der oben erwähnten Identifikations-unabhängigkeit selbst-referenzieller Bezugnahmen. Entscheidend ist ferner, dass IF bzw. die Identifikations-Unabhängigkeit von Ich-Gedanken wiederum an die Stelle von apriorischem Wissen über sich selbst tritt. Ein anderes Äquivalent von apriorischem Wissen über sich selbst oder gar über die Welt gibt es für Evans schlicht nicht. Freilich ist auch jene Form von Selbstkenntnis, die IF aufweist, nicht apriorisch im eigentlichen Sinn, insofern auch sie wesentlich auf externen Informationen beruht.

Um den Kontrast von IF zur Apriorizität im Zusammenhang mit EPS genauer zu sehen, ist es hilfreich, sich das Verhältnis zwischen *Immunität* auf der einen, und *Infallibilität* und *Inkorrigibilität* auf der anderen Seite klarzumachen. Selbstidentifikationen, die immun gegen Fehlidentifizierung sind, sind weder infallibel in Bezug auf die Selbstzuschreibungen, die Aussagen über die Zeit hinweg machen, und schon gar nicht in Bezug auf Identität der Person, die die Subjektstelle einnimmt, d. i. in Bezug auf personale Identität (VR, 196; 213ff.; 228f.).¹¹ Bestimmte Fragen danach, ob *ich* (noch) dieselbe Person bin wie vor 15 Jahren, oder ob ich morgen das glauben werde, was ich heute glaube, sind freilich sinnvoll, und bilden zum Teil den Reiz der metaphysischen Spekulationen über personale Identität. IF- oder Ich-Gedanken sind aber auch nicht inkorrigibel in Bezug auf allfällige zugeschriebenen körperlichen oder mentalen Eigenschaften oder den repräsentationalen *Gehalt* von Ich-Gedanken (das heißt in Bezug auf all dessen, was die Stelle von ϕ einnimmt). Sie sind lediglich inkorrigibel in Bezug auf die Identifizierung des Objekts/Subjekts (d. i., der Person) *zum Zeitpunkt* der Instanziierung eines Ich-Gedankens, dem jene Eigenschaften zugeschrieben werden: *Ich bin jetzt* ϕ , und das Subjekt, das *sich jetzt* ϕ zuschreibt, ist (jetzt) nicht etwa irgendjemand anderer. IF von Ich-Gedanken schließt also keineswegs aus und ist kompatibel mit der Möglichkeit, dass ein Subjekt das Objekt seiner demonstrativen Gedanken oder auch die körperlichen oder mentalen Prädikate, die er sich selbst zuschreibt, radikal fehlidentifizieren kann (vgl. VR, 179). Wie bei Shoemaker, der diese Eigenschaft von Selbstkenntnis als erster präzise formuliert hat (Shoemaker 1968), ist IF von selbstreferenziellen Ausdrücken *nicht* darauf zurückzuführen, dass Selbstzuschreibungen als solche oder hinsichtlich der zugeschriebenen Eigenschaften irrtumsresistent oder inkorrigibel sind, sondern darauf, dass das Subjekt, das sich etwas selbst zuschreibt, sich selbst gegenüber in keine Relation der Identifizierung tritt.

Mit dieser Konzeption von IF bezüglich Ich-Gedanken haben wir alles an der Hand, um nun Evans Kompatibilismus zu skizzieren.

4.2 Evans' alternativer Kompatibilismus

¹¹ Ich komme auf letzteres Problem im nächsten Abschnitt zurück.

Wie wir gesehen haben, scheint fast alles bei Evans zunächst geradehin in das inkompatibilistische Dilemma zu führen. So sind die Informationen, die uns die eigenen mentalen Zustände liefern, wenn wir introspektives Wissen von ihnen haben, nicht diesen Zuständen interne Informationen, sondern Informationen über die Welt und/oder über die (raum-zeitlich lokalisierbare) Personen, die sich in den betreffenden Informationszuständen befinden. Entsprechend legen die Subjekt-externen Faktoren bzw. die kausal-informationellen Links, die das Subjekt zu diesen hat, fest, *was* es weiß, wenn es introspektives Wissen von seinen eigenen mentalen Zuständen hat. Ferner müssen, gemäß Evans' russell-scher Konzeption von Gedanken, die Gegenstände bzw. Sachverhalte, die die mentalen (Glaubens-)Zustände individuierten, auch tatsächlich existieren. Ja, mehr noch: Wenn ein Subjekt sich in einem Informationszustand befindet, der einen kognitiven Gehalt über die Welt liefert, befindet es sich in einem Zustand, der ihm ipso facto all diejenigen Informationen liefert, die notwendig sind, um Aussagen über sich zu machen bzw. selbstreferenzielle Gedanken zu haben (VR, 230f.).

Doch Evans vertritt, zur Erinnerung, eben keinen Standard-Externalismus, und entsprechend sind die Individuationsbedingungen von mentalen Zuständen bei Evans eine komplexe multidimensionale Angelegenheit: *Qua Informationszustände* werden die mentalen Zustände eines Subjekts S durch S externe Entitäten (in der demonstrativ identifizierbaren Umwelt von S) individuiert. Diese bilden die Quelle von S's Informationszuständen, und zu ihnen steht S in bestimmten kausalen Relationen; *qua Ich-Gedanken* werden sie aber durch die jeweilige raum-zeitliche und körperliche Person P individuiert, die die Stelle von S in der objektiven Welt einnimmt, und die S selbst, wie wir gesehen haben, auf seiner egozentrischen Karte lokalisieren muss können (VR, 253). Nun ist aber das *Wissen* bzw. die jeweiligen Vorstellungen (*Ideas*), die S *über sich selbst* hat, wiederum abhängig a.) von dem Vermögen einer grundlegenden Selbstidentifikation, die immun gegen Fehlidentifizierung ist (und einer damit einhergehenden „grundlegenden Ich-Vorstellung“) und b.) von den (nicht-irrtumsresistenten und nicht-mental) Informationen, auf deren Grundlage S die *jeweiligen* Ich-Vorstellungen ausbildet und in eine Relation zu P setzt. Das heißt aber, dass S's Ich-Gedanken zwar nicht von irgendwelchen inneren mentalen Zuständen, sondern von externen Objekten und Sachverhalten handeln, gleichwohl sind sie zum Teil durch kriterienlose Selbstidentifikation konstituiert, und eben nicht allein durch externe Individuationskriterien. Sie sind aber auch, sofern S eben *Informationen* von den externen Entitäten beziehen muss können, über die seine Gedanken handeln (und nicht *nur* in *kausal*en Relationen dazu steht) – entgegen (AI) und aufgrund Russells Prinzip – *nicht unabhängig* von den *epistemischen* Individuationskriterien für Gedanken (cf. VR, 256f.).

Gemäß Evans' alternativem Kompatibilismus sind also selbstbewusste Gedanken *epistemisch* ausgezeichnet. Denn S hat gar keine selbstbewussten Gedanken, wenn es nicht eine – identifikations- bzw. kriterienlose *grundlegende* Konzeption von sich selbst als Subjekt seiner Gedanken hat.¹² Dieses Wissen ist

¹² Vgl. dazu auch Burges' (1988) ähnliches Konzept von *basic self-knowledge*.

jedoch eine Art Hintergrundwissen, und um dieses zu haben, muss S *nicht* die Individuationskriterien seiner Gedanken identifizieren und auch nicht die jeweiligen Individuationsbedingungen kennen, sondern lediglich dazu disponiert sein, sich selbst als raum-zeitlich lokalisiertes Subjekt kriterienlos zu identifizieren.

Als unmittelbare Antwort auf das mögliche inkompatibilistische Argument, das ich weiter oben konstruiert habe (s. Abschnitt 3), können wir nun folgendes kompatibilistische Argument konstruieren:

- (1) Immer wenn ein Subjekt S einen Ich-Gedanken (mit propositionalem Inhalt) hat, gibt es einen extern individuierten Informationszustand IZ, in dem sich S befindet.
- (2) Wenn S aufgrund von extern individuierten Informationen korrekt urteilt, dass *p*, hat S a priori Fehlidentifizierungs-immune Kenntnis davon, dass *er/sie glaubt*, dass *p*.
- (3) S kann aber aus der Proposition ‚Ich glaube, dass *p*‘ weder inferenziell noch apriorisch auf *p* schließen, noch dass sein Glaube, ‚dass *p*‘ gerechtfertigt ist.
- (4) Der Inkompatibilismus ist falsch.

S kann aber nicht nur *nicht* inferenziell oder gar a priori auf irgendwelchen externen Individuationskriterien seiner Gedanken schließen könnte, wie der Inkompatibilist behauptet – im Gegenteil: Ein Subjekt hat überhaupt nur Wissen darüber, was es für Gedanken hat, wenn es eine Konzeption der (externen) Objekte seiner Gedanken hat. Die Implikation ist für Evans sozusagen eine Einbahnstraße von der Welt zum Geist: Zwar muss tatsächlich ein Gegenstand G existieren, wenn man einen russellschen Gedanken hat, aber S kann durch das Haben eines G-Gedankens *allein* nicht wissen, ob G tatsächlich existiert, es könnte einfach der Fall sein, dass S *de facto* gar keinen G-Gedanken hat, und um *dies* festzustellen, muss S auf G, nicht auf den G-Gedanken blicken. Zu beachten ist ferner, dass das Subjekt in keinerlei epistemischen Relation zu IZ (in (1)) und auch nicht zu IZ's Individuationsbedingungen (in (2)), sondern über die relevanten Informationen verfügt, sofern es sich in einem solchen Informationszustand (qua mentalem Zustand) *befindet*:

His internal state cannot in any sense become an *object* to him. (He is *in* it). However, a subject can gain knowledge of his informational states in a very simply way: by re-using precisely those skills of conceptualization that he uses to make judgments about the world. [...] there is no *informational* state which stands to the internal state as that internal state to the state of the world. (VR, 227f.)

Wenn wir uns nun statt auf die epistemischen, vielmehr auf die referenziellen Aspekte inkompatibilistischer Argumente fokussieren, können wir eine Version von Evans' kompatibilistischem Argument auch so formulieren:

- (1) Wenn ein S einen russellschen Gedanken über G hat, dann gab/gibt es ein G und einen kausal-informationalen Link zwischen S und G.
- (2) S hat einen russellschen Gedanken über einen Gegenstand G.
- (3) S kann *nicht* durch bloße Introspektion, welche S Wissen über (1) und (2) lieferte, gepaart mit einer Modus-ponens-Inferenz, d. i., a priori, folgern:
 - (3a) ‚Es gab/gibt G und einen kausal-informationalen Link L zwischen S und G.‘
- (4) Der Inkompatibilismus ist falsch.

Wie bereits klar sein sollte, liegt der Grund für die negative These in (3) darin, dass es für Evans keine, wie auch immer geartete Form von Selbst-Kennntnis bezüglich (1) und (2) gibt, die verbürgen würde, dass die Aussage (3a) *wahr* ist bzw. dass es *tatsächlich* den betreffenden kausal-informationalen Link gibt. Im Gegenteil, erst das Vorliegen von L verbürgt, dass S einen G-Gedanken hat. Denn es gibt keinen Informationszustand *in* S, den S durch Introspektion eruieren könnte, und welcher S Kennntnis über das Vorliegen von L lieferte, jenseits genau desjenigen Informationszustandes, der durch L, und mithin das Vorliegen von G, konstituiert wird. Um herauszufinden, ob *dies* der Fall ist, muss S auf L und G blicken, nicht auf irgendwelche interne Zustände.

5. Einwände und Repliken

Ich habe bisher argumentiert, dass Evans einen Kompatibilismus anbietet, der die Skylla des Anti-Externalismus und die Charybdis der Leugnung der privilegierten Weise, wie wir Kennntnis von unseren eigenen Gedanken haben, umschiffen kann. Ich denke, dass Evans mit seinem Vorschlag auf der richtigen Fährte liegt. Allerdings möchte ich, auch auf die Gefahr hin, dass dies eine ungewöhnliche Dialektik in meiner Argumentation erzeugt, auf eminente Probleme in Evans' Lösungsvorschlag hinweisen – ohne, wohlgemerkt, dass ich selbst mit Evans hinreichend Argumente aufbringen werde können, diese Einwände restlos zu entkräften. Ich möchte zwei Gruppen von Einwänden hervorstreichen und beide dann mit einem Gedankenexperiment verschärfen.

Die erste Gruppe Einwänden betrifft die Transparenzthese und deren vermeintlich allzu weit reichende Radikalität. Fraglich ist hier welche epistemische Rolle externe Individuationsfaktoren bei *nicht-propositionalen* Formen von Selbstkennntnis spielen könnten. Welche externen Äquivalente gibt es, die man konsultieren könnte, um herauszufinden, welche nicht-propositionalen mentalen Zustände man hat, wie etwa phänomenale Zustände wie Schmerzen oder Stimmungen? (vgl. auch Campbell 2005)

Ein weiteres Problem im Zusammenhang mit der Transparenzthese betrifft die Verfügbarkeit und Identifizierung externer Kriterien von personaler Identität und des Materials für grundlegende Ich-Vorstellungen. Das Problem

lässt sich als ein zirkuläres Begründungsproblem ausdrücken: Gibt es irgendwelche externe Entitäten, die eine Quelle von Informationen sein könnten, auf die eine grundlegende Kenntnis von sich selbst beruht, ohne dass diese grundlegende Ich-Konzeption Teil dieser Information ist oder von dieser abgeleitet ist, zumal die grundlegende Ich-Vorstellung eben kein *Inhalt* von Ich-Gedanken oder Selbstzuschreibung ist, sondern diese ko-konstituieren (vgl. auch Rössler 1994; Bar-On 2004, 116, und Barz 2012, 26ff.).

Evans selbst scheinen diese beiden Probleme freilich nicht entgangen. So diskutiert er letzteres kurz in Zusammenhang mit dem Problem externer Kriterien für die Identifikation von sich selbst als über die Zeit hinweg identisches, oder personales *Selbst*: „How can it be that we can have knowledge of a state of affairs which involves a substantial and persisting self, simply by being aware of (still worse, by merely appearing to be aware of) a state of the world?“ (VR, 231) Dass es für Evans nicht hinlangt, (diachrone oder synchrone) personale Identität einfach mit dem Modell von Selbstidentifikation vorauszusetzen, oder personale Identität in die referenzielle Struktur Ich-Gedanken gleichsam analytisch hineinzuzinterpretieren, das wird an seiner Kritik an Anscombes Modell von Ich-Gedanken manifest (Anscombe 1975) (VR, 213f.). Evans' positiver Vorschlag ist aber weniger eindeutig, er involviert jedenfalls eine komplexe Diskussion von Erinnerung als eine *praktische Fähigkeit*, über die Zeit hinweg über die eigenen (Ich-)Gedanken gewissermaßen Buch zu führen oder im Auge zu behalten (VR, 235-248). Evans spricht an einer Stelle in Zusammenhang mit der Fähigkeit, den eigenen (räumlichen) Platz über kinästhetische Bewegungen hinweg zu orten, von einer „*skill*“ als „ability to keep track“ (VR, 236). Ähnlich könnten wir wohl die zeitliche Selbstverortung und Selbstidentifikation über diachrone und mentale bzw. psychologische Veränderung hinweg mit Evans als ein „praktisches Wissen“ (*practical knowledge*) (VR, 234), oder einer Fähigkeit und deren praktische Konsequenzen konstruieren. Diese wäre dann daran geknüpft, Unterschiede feststellen zu können zwischen dem Haben oder Nicht-Haben eines mentalen Zustandes und zwischen dem, dass *ich* sie habe oder hatte und dass jemand anderer sie hatte oder hat. Es wird eben auch einen praktischen und nicht nur einen theoretisch-epistemischen Unterschied machen, ob ich *glaube* oder *nicht glaube*, dass es morgen regnen wird, oder dass *ich* gestern Zahnschmerzen hatte, oder ob *du* Zahnschmerzen hattest. Im einen Fall werde ich morgen einen Regenschirm auf die Demo mitnehmen, oder heute zum Zahnarzt gehen, im anderen eben nicht.

Wenn Evans die Analogien und Unterschiede in der Weise notiert, in der jemand propositionale, perzeptuelle und phänomenale Tatsachen verifizieren kann, und gleichzeitig die radikale Transparenzthese in Bezug auch auf paradigmatische phänomenale Zustände wie Schmerzen nicht aufgeben mag, dann scheint er genau in diese Richtung zu deuten.

A subject's knowledge of what is in question when, for example, he is in pain, or when he sees a tree, can seem to be very similar to his knowledge of what is in question when an observable state of the world obtains. When an eclipse of the sun occurs, he might say, such-and-such experiences are to be expected; and he knows which they are—he can manifest this knowledge by his performance in suitable tests. Equally, he might say, when he is in pain, such-and-

such experiences are to be expected; and he knows which ones they are. There is certainly a difference between his being in pain and anyone else's being in pain; and this difference is one which he can detect, just as he can detect the difference between an eclipse of the sun and an eclipse of the moon. (VR, 233; vgl. VR, 208)

Die zweite Gruppe von Einwänden ist gravierender, zumal das Problem, das darin zum Ausdruck kommt, wohl kaum im Rahmen von Evans' Theorie von Selbstidentifikation lösbar ist; im Gegenteil, erst dieser Rahmen scheint das Problem virulent zu machen. Das Problem betrifft *negative* und *leere* Selbstidentifikationen. Erinnern wir uns: Jede Selbstidentifikation setzt für Evans folgendes voraus:

- (1) Wenn ein Subjekt S selbst-referenzielle Ich-Gedanken hat, dann gibt es eine Person, die sich immun gegen Fehlidentifizierung auf einer raum-zeitlichen egozentrischen kognitiven Karte lokalisieren kann, und diese Person ist identisch mit S.

Was passiert nun, wenn das Subjekt eine Art negative Selbstidentifikation vollzieht:

- (2) S denkt oder äußert den Ich-Gedanken:
 - (2a) *„Ich bin (oder war) keine Person, die auf einer raum-zeitlichen egozentrischen kognitiven Karte lokalisierbar ist (war).“*

Die Frage ist nun, ob (2a) durch S-unabhängige, extern individuierte Informationen, die S von sich selbst hat, widerlegt wird, oder aber schlicht dadurch, dass S einen Ich-Gedanken hat, und das heißt Gedanken des Typs, die (1) konstitutiv voraussetzen (vgl. auch Roos 2004). Diese Fragen lassen sich nicht ganz eindeutig beantworten. Denn, zum einen stellt sich die weitere Frage, was jene S-unabhängigen Informationen sein sollten, die (2a) falsifizieren könnten, wenn sie denn nicht Faktoren sind, die eben in (2a) konstitutiv vorausgesetzt sind (d. i., die Tatsache, dass das Subjekt und das Objekt der betreffenden Selbstidentifikation eine raum-zeitlich lokalisierbare Person ist). Aber heißt das nicht, dass (2a) schlicht leer oder selbstwidersprüchlich im Stile des Lügner-Paradoxes ist? Dies wiederum scheint intuitiv nicht ganz plausibel. Denn können wir uns nicht sehr wohl vorstellen, dass wir, zumindest teilweise, unsere Fähigkeit verlieren können, uns selbst auf ego-zentrischen kognitiven Karten zu lokalisieren, und gleichwohl selbstbewusste Gedanken mit intentionalem Gehalt haben bzw. referenziell gehaltvolle Aussagen tätigen können. Man denke hier etwa an partielle Amnesien.

Diese beiden Probleme der Individuation phänomenalen Selbstbewusstseins bzw. jener negativer und leerer Selbstidentifikationen lassen sich noch erheblich mit dem klassischen *Brain in the Vat* (BIV) Gedankenexperiment (Putnam 1981) verschärfen. Stellen wir uns nämlich folgende Variationen vor:

- (3a) Ein Gehirn im Tank (BIV) denkt, ‚Ich habe Fußschmerzen‘.

Und, um die Sache endgültig zu verkomplizieren, was passiert, wenn ein Gehirn im Tank den obigen Ich-Gedanken (2a) denkt, und mithin eine negative und ex hypothesi auch leeren Ich-Gedanken hat.

- (3b) Ein BIV denkt den Gedanken ‚Ich bin *keine* Person, die auf einer raum-zeitlichen egozentrischen kognitiven Karte lokalisierbar ist.‘

BIV's Selbstkonzeption basiert ex hypothesi auf keinen externen Informationen. BIV's (physikalische und evtl. mentale) Selbstzuschreibungen sind also referenziell gesehen leer. Denn ein BIV *formuliert* zwar insofern Ich-Gedanken, als es denken kann; doch sofern sein kognitives „Kontrollzentrum“ (VR, 254) quasi anderswo ist als der Gegenstand, den es in Ich-Gedanken identifiziert (der nämlich nirgendwo ist), schlägt seine grundlegende Selbstidentifikation fehl. Der Grund liegt darin, dass BIV's Ich-Gedanken einfach keine raum-zeitliche, körperliche Person herausgreifen, da ja BIV einen solchen Körper in der BIV-Welt ex hypothesi nicht hat, zumal seine Ich-Gedanken nicht in Handlungen manifestierbar sind.

Wir scheinen in einem Paradox gefangen: Wie kann nämlich ein BIV dann überhaupt eine grundlegende Ich-Konzeption haben, die es aber eben haben muss, um Ich-Gedanken überhaupt *formulieren* zu können? Gibt es aber nicht, so könnte eine mögliche Replik lauten, einen egozentrischen ‚Referenzrahmen‘ (*frame of reference*) (VR, 265) an Informationen und demonstrativer Identifikation auch in einer BIV-Welt? Es ist plausibel anzunehmen, dass dem so ist, und es scheint, dass auch Evans dieser Ansicht war – zumindest, wenn es um ähnliche Gedankenexperimente, die auf oft widerstreitende Intuitionen bezüglich entkörpernten oder körperlosen Gehirnen, oder auf mehrere Körper verteilten Gehirnen oder auf mehrere Gehirne verteilten Bewusstseins rekurrieren, wie jene von Dennett (1978) oder Nagel (1971) (VR, 253-255), oder um Selbstbezugnahmen von Doppelgängern in *Twin Earth*-Szenarien à la Putnam (1975) (VR, 263ff.).¹³ Evans scheint der Ansicht, dass wir auch in solchen Fällen eine Art raumzeitlich-lokalisierbaren referenziellen Rest hätten, der für die demonstrative und selbstbezüglich Identifikation ausreichte, und demzufolge das der egozentrische Rahmen eben ‚dort‘ ist, wo des Sprechers/Denkens *Gehirn* ist:

[...] the reason we do not find the ‚disembodied brain in a vat‘ case very disturbing, conceptually, is that the brain is also the last remaining part of the subject's body. (The case is often presented as a limiting case of amputation.) A tiny foothold is thus provided for the idea that the subject is where the brain is, and hence for the idea that the brain is what the subject is. (VR, 255)

¹³ Vgl. auch zu demonstrativen Bezugnahme und gegenständlicher Re-Identifikation in Zwillingserde-Szenarien, VR, 278-282. Ich habe das *Twin Earth*- bzw. BIV-Gedankenexperiment Putnams ausführlich in Szanto 2012 (267-285 bzw. 505-516) diskutiert.

Doch was, wenn überhaupt etwas, legt referenziellen Unterschied zwischen einem Ich-Gedanken, den ein BIV hat und einen, den ein Nicht-BIV hat, fest. Denn allein externe Informationen können es ja, wie wir gesehen, für Evans nicht sein. Ferner, sind *beide* immun gegen Fehlidentifizierung? Und wenn nicht, was verbürgt die Immunität der Ich-Gedanken von Nicht-BIVs im Unterschied zu BIVs? Es kann jedenfalls kein rein semantischer Unterschied sein, denn beiden Typen von Gedanken sind ihrer semantischen Struktur nach identisch, aber auch epistemologisch gesehen sind BIV- und Nicht-BIV-Gedanken ex hypothesi analog, sodass auch die epistemologische Struktur keinen Unterschied machen wird. Diese Fragen werden wohl, zumindest *innerhalb* des semantisch-referenziellen und epistemologischen Modells, den die Theorie der (Selbst-)Bezugnahmen Evans' vorgibt, offenbleiben müssen.

6. Schlussbemerkung

Ich habe argumentiert, dass Evans informationstheoretisch gewendete, radikal-externalistische Theorie von (Selbst-)Referenz, gepaart mit seinen nicht minder radikalen epistemologischen Thesen bezüglich der Transparenz des Geistes und der Immunität gegen Fehlidentifizierung von Ich-Gedanken einen originellen und überzeugenden Ausweg aus dem Dilemma anbietet, den die vermeintliche Inkompatibilität von Externalismus und Selbstkenntnis darstellt. Es scheint jedoch, als wären wir mit den eben diskutierten Einwänden nur erneut in einer Sackgasse gefangen. Doch wie auch immer man auf dieses beiden Typen von Einwänden antworten mag, und ob sie überhaupt zu parieren sind, entscheidend für das Ziel dieses Aufsatz ist, dass keine der beiden Evans Argumente für die *Kompatibilität* externalistischer und EPS-basierter Intuitionen selbst betrifft. Das heißt, auch wenn die Einwände als solche triftig genug wären, um zu zeigen, dass a.) die Transparenzthese zu radikal in Bezug auf (phänomenale) Selbstkenntnis ist, oder b.) dass das Modell IF-basierter Selbstidentifikation auf Grund Evans' sonstiger Hypothesen zur Lokalisierung psychophysischer Personen in Raum und Zeit und über die Zeit hinweg unlösbaren Problemen ausgesetzt ist – dies ändert nichts an einer möglichen, vielleicht etwas schwächeren, aber, angesichts der Komplexität des Problems, wie ich denke, immer noch durchaus attraktiven konditionalen Interpretation von Evans' Lösungsvorschlag: Demzufolge bietet Evans eine Lösung des besagten Dilemmas, *falls* wir seinen radikal-externalistischen Kernthesen bezüglich der Transparenz des Geistes und der Fehlidentifikations-Immunität von Selbstzuschreibung im Allgemeinen, oder auch nur in Bezug auf bestimmte Ich-Gedankens, etwa nur in Bezug auf *propositionale* Ich-Gedanken, akzeptieren können. Dies, so hoffe ich gezeigt zu haben, ist angesichts der Dringlichkeit des Problems durchaus nicht geringzuschätzen, zumal wir, wie wir gesehen haben, bei den Opponenten weit inakzeptablere Prämissen, wie etwa die Verdoppelung des propositionalen Wissens, akzeptieren müssen, damit ihre Reductio-Argumente überhaupt Boden unter den Füßen bekommen.

Referenzen

- Anscombe, G.E.M. (1975): The first person. In: *Mind and Language*, hg. von In: S. Guttenplan, Oxford, S. 45–65.
- Bar-On, D. (2004): Externalism and Self-Knowledge: Content, Use, and Expression. In: *Noûs* 38 (3), S. 430–455.
- Barz, W. (2012): *Die Transparenz des Geistes*. Berlin.
- Bermúdez, J.L. (2005): Evans and the Sense of “I”. In: *Thought, Reference, and Experience: Themes from the Philosophy of Gareth Evans*, hg. von J.L. Bermúdez. Oxford, S. 164–194.
- Bilgrami, A. (1992): Can Externalism Be Reconciled with Self-Knowledge? In: *Philosophical Topics* 20 (1), S. 233–267.
- Brown, J. (1995): The incompatibility of anti-individualism and privileged access. In: *Analysis* 55 (3), S. 149–156.
- (1999): Boghossian on externalism and privileged access. In: *Analysis* 59 (1), S. 52–59.
- (2004): *Anti-Individualism and Knowledge*. Cambridge, MA/London.
- (2007): Externalism in Mind and Epistemology. In: *Internalism and Externalism in Semantics and Epistemology*, hg. von S. Goldberg. Oxford, S. 13–34.
- Brueckner, A.L. (1992): What an Anti-Individualist Knows A Priori. In: *Analysis* 52 (2), S. 111–118.
- (2007): Externalism and Privileged Access Are Consistent. In: *Contemporary Debates in Philosophy of Mind*, hg. von B. P. McLaughlin & J. D. Cohen, S. 37–51.
- Burge, T. (1979): Individualism and the Mental. In: *Midwest Studies in Philosophy* 4 (1), S. 73–122.
- (1986): Individualism and Psychology. In: *The Philosophical Review* 95 (1), S. 3–45.
- (1988): Individualism and Self-Knowledge. In: *Journal of Philosophy* 85 (11), S. 649–663.
- Campbell, J. (2005): Information Processing, Phenomenal Consciousness, and Molyneux’s Question. In: *Thought, Reference, and Experience: Themes from the Philosophy of Gareth Evans*, hg. von J.L. Bermúdez. Oxford, S. 195–219.
- Castañeda, H.-N. (1966): “He”: On the Logic of Self-Consciousness. In: *Ratio* 8, S. 130–157.
- Chisholm, R.M. (1981): *The First Person: An Essay on Reference and Intentionality*. Minneapolis.
- Davidson (1984): First person authority. In: *Dialectica* 38 (2–3), S. 101–111.
- (1987): Knowing one’s own mind. In: *Proceedings and addresses of the American philosophical association* 60 (3), S. 441–458.
- (1988): Reply to Burge. In: *The Journal of Philosophy* 85 (11), S. 664–665.
- (1989): What is Present to the Mind? In: *Philosophical Issues* 1, S. 197–213.
- Dennett, D.C. (1978). In: D.C. Dennett: *Brainstorms: Philosophical Essays on Mind and Psychology*. Brighton, S. 310–323.
- Dretske, F. (1995): *Naturalizing the Mind*. Cambridge, MA/London.
- Evans, G. (1973): The Causal Theory of Names. In: *Proceedings of the Aristotelian Society, Suppl. Vols.* 47, S. 187–208.
- (1980): Commentary on Jerry A. Fodor’s “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology.” In: *The Behavioral and Brain Sciences* 3, S. 79–80.

- (1982) [=VR]: *The Varieties of Reference*, hg. von J. McDowell. Oxford.
- Falvey, K. (2000): The Compatibility of Anti-Individualism and Privileged Access. In: *Analysis* 60 (1), S. 137–142.
- Falvey, K. & Owens, J. (1994): Externalism, Self-Knowledge, and Scepticism. In: *The Philosophical Review* 103 (1), S. 107–137.
- Goldberg, S. (2015): Introduction. In: *Externalism, Self-Knowledge, and Skepticism: New Essays*, hg. von S. C. Goldberg. Cambridge.
- Heil, J. (1988): Privileged Access. In: *Mind* 157 (386), S. 238–251.
- Jacob, P. (2004): Do We Know how We Know Our Own Minds yet? In: *The Externalist Challenge*, hg. von R. Schantz. Berlin/New York, S. 401–418.
- Kallestrup, J. (2011): Recent Work on McKinsey’s Paradox. In: *Analysis* 71 (1), S. 157–171.
- Kripke, S.A. (1972): Naming and necessity. In: *Semantics of Natural Language*, hg. von D. Davidson & G. Harman. Dordrecht, S. 253–355.
- (2008): Frege’s Theory of Sense and Reference: Some Exegetical Notes. In: *Theoria* 74 (3), S. 181–218.
- Lewis, D. (1979): Attitudes De Dicto and De Se. In: *The Philosophical Review* 88 (4), S. 513–543.
- Ludlow, P. (1995): Externalism, Self-Knowledge, and the Prevalence of Slow Switching. In: *Analysis* 55 (1), S. 45–49.
- Ludlow, P. & Martin, N. (1998): „Introduction.“ In: *Externalism and Self-Knowledge*, hg. von P. Ludlow & N. Martin. Stanford.
- Macdonald, C. (1998): Externalism and Authoritative Self-Knowledge. In: *Knowing Our Own Minds*, hg. von C. Wright, B. Smith, & C. Macdonald. Oxford, S. 123–155.
- McLaughlin, B.P. & Tye, M. (1998). Externalism, Twin Earth, and Self-Knowledge. In: *Knowing Our Own Minds*, hg. von C. Wright, B. Smith, & C. Macdonald. Oxford, S. 285–320.
- McKinsey, M. (1991): Anti-Individualism and Privileged Access. In: *Analysis* 51 (1), S. 9–16.
- (2007): Externalism and Privileged Access Are Inconsistent. In: *Contemporary Debates in Philosophy of Mind*, hg. von B. P. McLaughlin, & J. D. Cohen. London, S. 52–63.
- Miller, R.W. (1997): Externalist self-knowledge and the scope of the a priori. In: *Analysis* 57 (1), S. 67–75.
- Nagel, T. (1971). Brain bisection and the unity of consciousness. In: *Synthese* 22 (3), S. 396–413.
- Nuccetelli, S. (2003): Introduction. In: *New Essays on Semantic Externalism and Self-Knowledge*, hg. von S. Nuccetelli. Cambridge, MA/London, S. 1–21.
- O’Brien, L. (1995): Evans on Self-Identification. In: *Noûs*, 29 (2), S. 232–247.
- Perry, J. (1979): The Problem of the Essential Indexical. In: *Noûs* 13 (1), S. 3–21.
- Putnam, H. (1975): The meaning of “meaning”. In: *Minnesota Studies in the Philosophy of Science* 7, S. 131–193.
- (1981): *Reason, Truth and History*. Cambridge.
- Roos, A. (2004): An Objection to Gareth Evans’ Account of Self-Identity. In: *Ratio* 17 (2), S. 207–217.
- Rorty, R. (1970): Incorrigibility as the Mark of the Mental. In: *Journal of Philosophy* 67 (12), S. 399–424.

- Rössler, Johannes (1994). Einleitung zu Evans, Gareth. In: *Analytische Theorien des Selbstbewußtseins*, hg. von M. Frank. Frankfurt a.M., S. 483–499.
- Rowlands, M. (2003): *Externalism: Putting Mind and World Back Together Again*. Montreal & Ithaca.
- Sawyer, S. (1998): Privileged access to the world. In: *Australasian Journal of Philosophy* 76 (4), 523–533.
- Shoemaker, S. (1968): Self-Reference and Self-Awareness. In: *The Journal of Philosophy* 65 (19), S. 555–567.
- Strawson, P.F. (1959): *Individuals: An Essay in Descriptive Metaphysics*. London.
- Szanto, T. (2012): *Bewusstsein, Intentionalität und Mentale Repräsentation: Husserl und die analytische Philosophie des Geistes*. Berlin/Boston.
- Warfield, T. (1992). Privileged self-knowledge and externalism are compatible. In: *Analysis* 52 (4), S. 232–237.
- Wright, C., Smith, B.C., & Macdonald, C. (1998): Introduction. In: *Knowing Our Own Minds*, hg. von C. Wright, B.C. Smith, & C. Macdonald. Oxford, S. 1–12.